DOI: 10.51790/2712-9942-2021-2-3-3

НЕКОТОРЫЕ МАТЕМАТИЧЕСКИЕ АСПЕКТЫ ПРОБЛЕМЫ ПОСТРОЕНИЯ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

В. Б. Бетелин[1], **В. А. Галкин**[2,3], **А. О. Дубовик**[2,3]

[1] *Федеральное государственное учреждение «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук», г. Москва, Российская Федерация, betelin@niisi.msk.ru*

[2] *Сургутский филиал Федерального государственного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук», г. Сургут, Российская Федерация, val-gal@yandex.ru*

[3] *Сургутский государственный университет, г. Сургут, Российская Федерация, alldubovik@gmail.com*

Аннотация: искусственные нейронные сети (ИНС) в настоящее время являются полем интенсивных исследований. Они зарекомендовали себя при решении задач распознавания образов, аудио и текстовой информации. Планируется их применение в медицине, в беспилотных автомобилях и летательных аппаратах. Однако крайне мало научных работ посвящено обсуждению возможности построения искусственного интеллекта (ИИ), способного эффективно решать очерченный круг задач. Отсутствует гарантия штатного функционирования ИИ в любой реальной, а не специально созданной ситуации.

В данной работе предпринимается попытка обоснования ненадежности функционирования современных искусственных нейронных сетей. Показывается, что задача построения интерполяционных многочленов является прообразом проблем, возникающих при создании ИНС. Известны примеры К.Д.Т. Рунге, С.Н. Бернштейна и общая теорема Фабера о том, что для любого наперед заданного натурального числа, соответствующего количеству узлов в интерполяционной таблице, найдется точка из области интерполяции и непрерывная функция, что интерполяционный многочлен не сходится к значению функции в этой точке при неограниченном росте числа узлов. Отсюда следует невозможность обеспечения эффективной работы ИИ лишь за счет неограниченного роста числа нейронов и объемов данных (Big Data), используемых в качестве обучающих выборок.

Ключевые слова: искусственные нейронные сети, проблемы применения нейронных сетей.

Благодарности: работа выполнена в рамках государственного задания ФГУ ФНЦ НИИСИ РАН (проведение фундаментальных научных исследований (47 ГП) по теме № 0065-2019-0007 «36.20 Развитие методов математического моделирования распределенных систем и соответствующих методов вычисления» (№ АААА-А19-119011590093-3).

Для цитирования: Бетелин В. Б., Галкин В. А., Дубовик А. О. Некоторые математические аспекты проблемы построения искусственных нейронных сетей. *Успехи кибернетики*. 2021;2(3):19–22. DOI: 10.51790/2712-9942-2021-2-3-3.

SOME MATHEMATICAL ASPECTS OF CONSTRUCTING ARTIFICIAL NEURAL NETWORKS

V. B. Betelin[1], **V. A. Galkin**[2], **A. O. Dubovik**[3]

[1] *Federal State Institution "Scientific Research Institute for System Analysis of the Russian Academy of Sciences", Moscow, Russian Federation, betelin@niisi.msk.ru*

[2] *Surgut Branch of Federal State Institute "Scientific Research Institute for System Analysis of the Russian Academy of Sciences", Surgut, Russian Federation, val-gal@yandex.ru*

[3] *Surgut State University, Surgut, Russian Federation, alldubovik@gmail.com*

Abstract: artificial neural networks (ANN) are currently a field of intensive research. They are a proven pattern/audio/text recognition tool. ANNs will be used in medicine, autonomous vehicles, and drones. Still, very few works discuss building artificial intelligence (AI) that can effectively solve the mentioned problems. There is no guarantee that AI will operate properly in any real-life, not simulated situation.

In this work, an attempt is made to prove the unreliability of modern artificial neural networks. It is shown that constructing interpolation polynomials is a prototype of the problems associated with the ANN generation. There are examples by C.D.T. Runge, S.N. Bernstein, and the general Faber theorem stating that for any predetermined natural number corresponding to the number of nodes in the lookup table there is a point from the interpolation region and a continuous function that the interpolation polynomial does not converge to the value of the function at this point as the number of nodes increases indefinitely. This means the impossibility of ensuring efficient AI operation only by an unlimited increase in the number of neurons and data volumes (Big Data) used as training datasets.

Keywords: artificial neural networks, neural network applications.

Acknowledgements: this study is the 47 GP government order contracted to the Scientific Research Institute for System Analysis, Russian Academy of Sciences, project No. 0065-2019-0007 36.20 Advancing Distribution System Simulation and Computation Methods (No. AAAA-A19-119011590093-3).

Cite this article: Betelin V. B., Galkin V. A., Dubovik A. O. Some Mathematical Aspects of Constructing Artificial Neural Networks. *Russian Journal of Cybernetics*. 2021;2(3):19–22. DOI: 10.51790/2712-9942-2021-2-3-3.

Во многочисленных вычислительных экспериментах по распознаванию образов, обработке акустической, видео и текстовой информации были найдены подходы к созданию программного обеспечения для практического решения ряда трудно формализуемых задач. Эти успехи породили огромный поток работ и интерес к созданию полуэмпирических методов, носящих название искусственные нейронные сети (ИНС).

Общий подход, лежащий в основе построения ИНС, состоит в принятии гипотезы возможности создания устройства, которое можно обучить на серии примеров принятию решений.

Настройка ИНС (аналог процедуры автоматизированного программирования структуры ИНС) состоит в подборе семейства искусственных нейронов (преобразований) и коэффициентов умножения вектора входного сигнала на основе дополнительно декларируемых принципов (например, оптимизации). Целью создания ИНС является в некотором смысле «оптимальная реконструкция» неизвестного отношения R на основе заданного «обучающего» набора $\tilde{R}$. По своей природе такая постановка задачи является некорректной ввиду существенной множественности ее решений («школьная» задача об отыскании кривой, соединяющей заданный набор точек). Таким образом, необходимым элементом построения ИНС служит введение целевого правила (критерия, функционала и т. п.) F, на основе применения которого выполняется построение «наилучшего» продолжения заданного отношения $\tilde{R}$ до отношения R_F, которое аппроксимирует R (например, посредством минимизации функционала F на некотором множестве параметров). Назовем R_F математической моделью реализации ИНС. Описанная схема обычно применяется для построения продолжения функций, в частности, в задачах интерполяции.

Будем далее рассматривать задачи построения интерполяционных многочленов как прообраз математических проблем, возникающих при создании ИНС. Известно, что при неограниченном наращивании объема обучающей совокупности R возникают проблемы с точностью ИНС. Здесь показательными являются примеры К. Д. Т. Рунге [1] и С. Н. Бернштейна [2] для интерполяционных многочленов с расходящейся последовательностью аппроксимаций при неограниченном увеличении количества узлов интерполяции, что для ИНС означает неограниченный рост числа нейронов. Теорема Фабера [3] гласит, что для любой интерполяционной таблицы размерности N найдутся непрерывная функция и точка на отрезке аппроксимации, такие, что интерполяционные многочлены не сходятся к указанной функции в этой точке при $N \to \infty$.

Примеры Рунге, Бернштейна и общий результат Фабера указывают на возможность глубокого заблуждения, что неограниченное наращивание нейронов в ИНС и большие объемы данных (Big Data), положенные в основу обучающих наборов, могут обеспечить ее эффективность в рамках ИИ. То же относится ко многим утверждениям о построении так называемых «цифровых двойников» за счет безграничного наращивания обучающих наборов данных.

Указанные замечания относятся к общей фундаментальной проблеме анализа вычислительной устойчивости ИНС при наращивании объемов обучающих наборов и, как следствие, — границ реализуемости и безопасной применимости систем ИИ, построенных на комплексах, взаимодействующих с ИНС.

Одна из первых связей между искусственным интеллектом и теоремой Геделя о неполноте была установлена британским философом Дж. Лукасом.

Любая формализованная арифметика состоит из языка (алфавита), аксиом (исходных истинных утверждений) и правил вывода, позволяющих из аксиом (и гипотез) получать новые истинные утверждения. Теорема Геделя о неполноте формализованной арифметики в несколько более вольной интерпретации, чем описано в [5], утверждает, что если формальная система (формализованная арифметика) непротиворечива, то существует выражение, формулируемое на языке этой системы, такая, что не выводима ни сама эта формула, ни ее отрицание. Т. е. истинность или ложность такого утверждения невозможно установить в рамках данной формальной системы.

Дж. Лукас изложил свои идеи в статье «Умы, машины и Гедель» [6]. Его идея гласит, что многие машины, которые являются электронными или вычислительными, включая компьютеры, представляют собой пример формальной системы. Кибернетические машины, такие как современный компьютер, имеют конечные операции и соответствуют определению формальной системы Геделя. Все операции компьютеров могут быть представлены в виде формул, и существуют правила вывода, которые задаются компьютерными алгоритмами.

Одной из целей исследований ИИ является достижение «сильного искусственного интеллекта», то есть общего ИИ человеческого уровня. В настоящее время ИИ строится как алгоритмы в машинах Тьюринга, которые являются последовательными аксиоматическими системами и поэтому подчиняются теореме Геделя.

Британские математик и физик Р. Пенроуз [7, 8] и философ Дж. Лукас [6] утверждают, что человеческое сознание превосходит машины Тьюринга, потому что человеческий разум посредством интроспекции может распознавать свои собственные несоответствия, что согласно теореме Геделя невозможно для машин Тьюринга. И поэтому для машин Тьюринга недоступна возможность воспроизводить черты человеческого разума, например, такую, как математическая проницательность. Из теоремы Левенгейма–Сколема следует, что математическую реальность невозможно однозначно включить в аксиоматическую систему [9], а значит, компьютер никогда не сможет вывести все теоремы.

Согласно теореме Геделя в любой непротиворечивой формальной системе на языке этой системы можно сформулировать утверждение, не выводимое вместе со своим отрицанием, т. е. формальная система является неполной. Для разрешения неопределенности необходимо добавить это утверждение или его отрицание в состав аксиом рассматриваемой формальной системы. Для машины Тьюринга это означает наличие внешней по отношению к ней более интеллектуально развитой системы, которой каждый раз необходимо выбирать, какую новую аксиому добавить в формальную систему и в соответствии с этим выбором исправить алгоритм работы машины Тьюринга. Такие интеллектуальные системы могут образовывать иерархию все более сложных и всеобъемлющих систем, на вершине которой находится человек. Необходимость такой иерархии следует из теоремы Левенгейма–Сколема, которая утверждает, что любая система аксиом допускает намного больше существенно различных интерпретаций, чем предполагалось при ее создании. Аксиомы не устанавливают пределов для интерпретаций или моделей, и поэтому необходимо существование корректора — человека. Аксиоматические системы, к которым применима теорема Левенгейма–Сколема, предназначаются для задания одной вполне конкретной интерпретации, и, будучи применимыми к совершенно различным моделям, они тем самым не соответствуют своему назначению.

Подводя итог сказанному, можно отметить, что прорыв, произведенный НС и методами глубокого обучения, обеспечил решение задач, отражающих периферийные функции человека. Сам процесс обдумывания, планирования своих действий («поведенческое» управление), логического «рассуждения» и абстрактного мышления пока не воспроизводится работой искусственных нейронных сетей. Эти процессы так или иначе будут воспроизводиться на основе объединения различных методов ИИ (не только и не столько нейросетевых) с имитационным моделированием.

ЛИТЕРАТУРА

1. Runge K. Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten. *Zeitschrift für Matematik und Physik*. 1901;46:224–243.
2. Бернштейн С. Н. Собрание сочинений. Т. 1–4. М., 1952–1964.
3. Дзядык В. К. *Введение в теорию равномерного приближения функций полиномами*. М.: Наука; 1977.
4. Тихонов А. Н., Арсенин В. Я. *Методы решения некорректных задач*. М.: Наука; 1979.

5. Новиков П. С. *Конструктивная логика с точки зрения классической*. М.: Наука; 1977. 328 с.
6. Lucas J. R. Minds, Machines and Gödel. *Philosophy*. 2009;36(137):112–127.
7. Penrose R. *The Emperor's New Mind*. Oxford University Press; 1989. 480 p.
8. Penrose R. *Shadows of the Mind*. Oxford University Press; 1994. 457 p.
9. Клайн М. *Математика. Утрата неопределенности*. М.: Мир; 1984. 434 с.